

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МІСЬКОГО ГОСПОДАРСТВА імені О. М. БЕКЕТОВА**

МЕТОДИЧНІ РЕКОМЕНДАЦІЇ

до самостійної роботи з дисциплін

**«ЕЛЕКТРОННИЙ ДОКУМЕНТООБІГ В ПРОЕКТУВАЛЬНИЙ
ДІЯЛЬНОСТІ», «ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В НАУКОВІЙ ТА
ПРЕЗЕНТАЦІЙНІЙ РОБОТІ», «ІНФОРМАЦІЙНІ СИСТЕМИ В
МЕНЕДЖМЕНТІ»**

(для студентів усіх форм навчання спеціальності 073 – Менеджмент)

**Харків
ХНУМГ ім. О. М. Бекетова
2019**

Методичні рекомендації до самостійної роботи з дисциплін «Електронний документообіг в проєктувальній діяльності», «Інформаційні технології в науковій та презентаційній роботі», «Інформаційні системи в менеджменті» (для студентів усіх форм навчання спеціальності 073 – Менеджмент) / Харків. нац. ун-т міськ. госп-ва ім. О. М. Бекетова ; уклад. : М. Ю. Карпенко, І. О. Гавриленко. – Харків : ХНУМГ ім. О. М. Бекетова, 2019. – 22 с.

Укладачі: М. Ю. Карпенко,
І. О. Гавриленко

Рецензент: О. Б. Костенко, канд. фіз.-мат. наук, доцент Харківського національного університету міського господарства імені О. М. Бекетова.

Рекомендовано кафедрою прикладної математики і інформаційних технологій, протокол № 1 від 31.08.2016.

ЗМІСТ

1 Вступ	4
2 Що таке сканування й розпізнавання	4
3 Які зображення можна розпізнати	5
4 Як розпізнати документ у програмі FineReader	6
4.1 Інтерфейс програми FineReader	6
4.2 Послідовність дій щодо розпізнавання документа	7
4.3 Завантаження зображень	8
4.4 Автокоригування зображень	8
4.5 Розмітка документа	10
4.6 Розпізнаємо зображення	12
4.7 Зберігаємо результат	12
4.8 Імпортуємо результат	13
4.9 Розпізнавання таблиці складної структури	13
5 Коригування зображень перед їхнім розпізнаванням	15
5.1 Як перезняти зображення	15
5.2 Штучний спалах та його використання	15
5.3 Послідовність дій з коригування зображень	16
5.3.1 Завантажуємо файл	17
5.3.2 Коригуємо розмір зображення	17
5.3.3 Прибираємо зайві деталі	18
5.3.4 Коригуємо геометрію	19
5.3.5 Коригуємо яскравість	19
6 Завдання до виконання роботи	21
Список рекомендованої літератури	22

1 Вступ

Опанувати роботу з програмою FineReader. Навчитися розпізнавати документ за допомогою програми FineReader. Одержати навички щодо коригування зображень засобами графічного редактора PhotoShop для їх подальшого розпізнавання.

2 Що таке сканування й розпізнавання

Сканування – це процес одержання графічної копії оригінального зображення. Сканувати документ можна по-різному. Наприклад, – за допомогою сканера або фотоапарата. І в тому і в іншому випадках результатом буде растровий графічний файл. Цей файл може бути кольоровим або чорно-білим, мати різну роздільність, може бути представлений у різних форматах (найчастіше «*.jpg» або «*.tif»). Важливо розуміти головне: растровий файл – це набір точок. Те, що на екрані або друкованому аркуші він виглядає як текст, ще ні про що не свідчить. Обробити цей «текст» програмою Word або Excel неможливо. Усе, що ви зможете з ним зробити – це вставити його у вигляді готового об'єкта. Відкоригувати растровий файл можна лише у спеціальній програмі – графічному редакторі. При цьому працювати з ним можна як з малюнком, а не як з текстовим документом.

Зайвий раз пояснимо різницю між текстом й сканованим документом. Уявіть, що ви вводите документ до програми Word, набираючи його на клавіатурі. Щоразу, натискаючи клавішу, до комп'ютера потрапляє її код. Фактично цей код і зберігається у «вордовському» документі. У цьому випадку для кожної букви документа відомо її походження. Наприклад, перший символ цього абзацу має код «0194», це буква «В», другий – код «0224», це символ «а» і т.д. При друкуванні або виведенні тексту на екран виконується ВІЗУАЛІЗАЦІЯ документа. Тобто виходячи з кодів, що є в тексті Word формує їх графічне зображення, іншими словами, «промальовує» букви. Коли ви коригуєте документ у програмі Word, ви змінюєте коди символів у документі. Саме тому одне натискання клавіші може змінити лише один символ. У графічному зображенні цього немає. Сканер нічого не друкує. Він вводить піксели (крапки) зображення. Будучі розташовані у певному порядку, ці піксели виглядають як текст, і не більше того. Той факт, що якась частина цього зображення – це символ «В», якась частина – символ «а» комп'ютеру невідома. Для нього ці символи – усього лише вибудований певним чином набір пікселів. Тому й відредагувати ці піксели програмою Word не можна.

Для того щоб перетворити графічний файл до текстового формату, його потрібно розпізнати. Цим займаються спеціальні програми оптичного розпізнавання образів (OCR). Однією з найпоширеніших програм цього класу є FineReader. Незважаючи на відмінності у деталях, мета у цих програм одна. Вони аналізують зображення, намагаючись знайти фрагменти із якимось змістом. Це можуть бути символи тексту, таблиці, малюнки. Тобто програма розпізнавання намагається з'ясувати структуру документа: знайти букви, слова, речення, орієнтуючись на специфіку їх зовнішнього вигляду. Потім вона формує текстовий документ, перевіряє його за системою словників, граматичних правил й пропонує користувачеві для остаточного контролю. Коли перевірку завершено, результат роботи програми розпізнавання можна зберегти у формат текстового процесора Word або електронної таблиці Excel. І тільки після цього ви одержите «нормальний» документ, з яким можете робити все, що завгодно.

Звичайно, на практиці не все так просто. Алгоритм розпізнавання – процес складний, у випадку неякісних зображень програма розпізнавання припуститься чисельних помилок. Щоб мінімізувати їх кількість, треба чітко уявляти, які мають бути зображення для доброго розпізнавання, як їх можна виправити у графічному редакторі, яким чином треба організувати сам процес розпізнавання.

3 Які зображення можна розпізнати

Для успішного розпізнавання зображення має відповідати певним вимогам, а саме:

- графічний файл повинен мати роздільність приблизно 300 dpi;
- його лінійний розмір при цій роздільності має бути близький до оригіналу. Навряд чи вам вдасться розпізнати сторінку документа з роздільністю 300 dpi розміром з поштову марку;
- зображення має бути чорно-білим або у відтінках сірого.

Це, так би мовити, технічні вимоги до оригіналів. Зрозуміло, є також інші, не менш важливі, умови для виконання цієї роботи. Наприклад, зображення має бути досить контрастним, букви – з чітко окресленими краями. Любе «сміття» на сканованому зображенні у вигляді позначок олівцем, плям, дефектів паперу програма розпізнавання спробує ідентифікувати як символи. Ці обставини треба враховувати при підготовці зображень для подальшого розпізнавання.

4 Як розпізнати документ у програмі FineReader

Спочатку навчимося розпізнавати якісні зображення. Тобто такі, що не потребують обробки. Такі зображення можуть бути отримані зі сканера, або внаслідок копіювання фрагментів екрана. Для прикладу ми візьмемо для розпізнавання фрагмент тексту з таблицею, що зображений на рисунку 1.

Какие изображения можно распознать		
Чтобы процесс распознавания был успешным, графическое изображение должно соответствовать определенным требованиям, а именно.		
Параметр	Значение	Комментарий
Разрешение	300 dpi	
Формат	«Jpg», «Tif», «Bmp»	Практически любой растровый формат
Тип	Цветное, оттенки серого, черно-белое	

Рисунок 1

Для розпізнавання будемо використати програму «FineReader 11.0» компанії ABBYY (<http://www.abbyy.ua/>), скачавши демонстраційну версію з <http://www.abbyy.ru/Download/>.

4.1 Інтерфейс програми FineReader

Вікно програми зображено на рисунку 2. У ньому представлено меню, під ним – панелі інструментів. Нижче панелі інструментів знаходяться п'ять великих іконок для керування програмою, а саме: «Scan&Read», «Открыть», «Распознать», «Проверить», «Microsoft Word». Основна частина вікна розбита на чотири області. Перелічимо їх у порядку зліва-направо й зверху-вниз. У лівій частині вікна міститься область сторінок документа. Кожна сторінка – це окреме зображення. Праворуч від області сторінок розташоване віконце «Изображение», де є фрагмент відсканованого документа та його розмітка (про неї поговоримо трохи пізніше).

У правій частині вікна показано результат розпізнавання. Це – текстовий документ, його можна правити прямо у програмі FineReader. У нижній частині вікна розташована область зі збільшеною копією зображення.

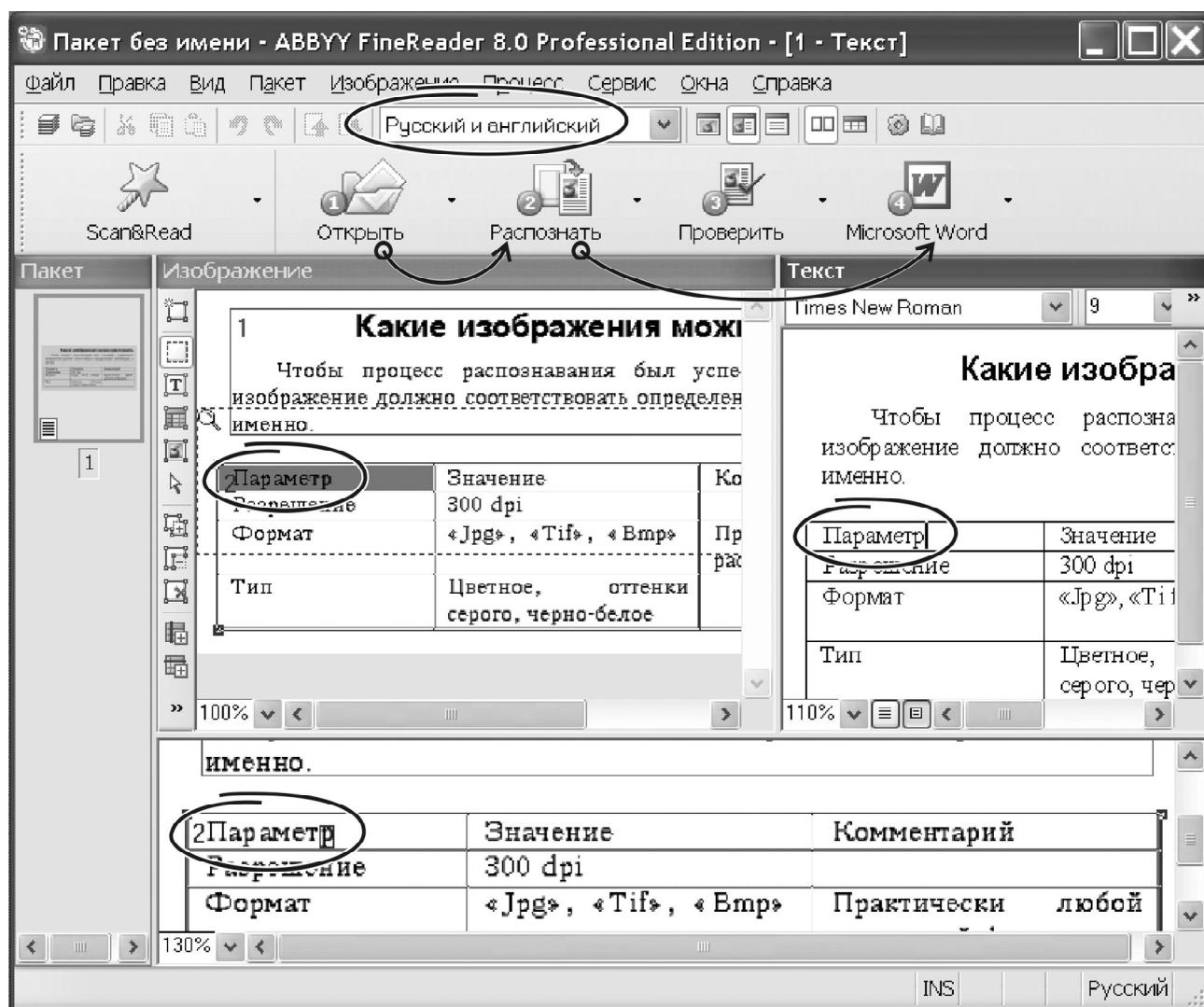


Рисунок 2

Між областю сторінок й вікном «Изображение» знаходиться робоче меню для розмітки зображення.

4.2 Послідовність дій щодо розпізнавання документа

Щоб розпізнати сканований документ, потрібно зробити так:

1. Завантажити одне або декілька зображень до програми FineReader.
2. Відредагувати зображення, щоб очистити його від «сміття», усунути перекручування рядків й т. і. Цей етап потрібен не завжди.
3. Розмітити зображення, тобто вказати, Як FineReader повинен інтерпретувати окремі ділянки зображення. Звичайно така операція потрібна

при складному форматуванні документа. Для простих випадком операцію розмітки FineReader прекрасно виконує в автоматичному режимі. 4. Розпізнати документ.

5. Переглянути результат роботи. Якщо FineReader помилився, можна відредагувати документ або поправити розмітку й повторити розпізнавання.

6. Імпортувати результат розпізнавання до зовнішнього додатку, наприклад, до програми Word або Excel.

4.3 Завантаження зображень

У нашому випадку дані для розпізнавання перебувають на жорсткому диску у вигляді графічних файлів. Натискаємо кнопку «Открыть» (на рис. 2 вона зображена з номером «1»). З'явиться вікно Провідника. У ньому знаходимо папку, виділяємо потрібні файли й клацаємо на кнопці «Открыть» або натискаємо «Enter». Завантажені зображення з'являться в області сторінок.

Для завантаження зображень нажміть комбінацію «Ctrl+O» (англійська буква «O»). Також можна скористатись меню «Файл → Открыть».

Кнопка «Scan&Read» служить для початку сканування. Якщо ви користуєтесь сканером, включіть його, покладіть документ й натисніть «Scan&Read». Подальша робота залежить від моделі сканера. Через кнопку «Scan&Read» можна також завантажити документ. Клацніть на маленькому трикутнику в лівій частині іконки «Scan&Read». Розкриється список можливих дій. Серед них буде операція «Открыть и распознать». Розкривши список дій кнопки «Открыть», ви побачите рядок «Сканировать изображение... Ctrl+K»... Таким чином, кнопки «Scan&Read» й «Открыть» багато у чому взаємозаміняють одна одну.

4.4 Автокоригування зображень

Операції коригування зображення зосереджені в меню «Изображение». Розберемо ті можливості, які доводиться використати найчастіше.

ПОВОРОТ ЗОБРАЖЕННЯ. Текст при розпізнаванні повинен бути зорієнтований «зверху вниз» і впливати «зліва направо». Якщо сторінка документа сканована з поворотом, її прийде переорієнтувати. Для цього ввійдіть в «Изображение → Повернуть/Отразить изображение». З'явиться

меню з п'яти пунктів: «Повернуть по часовой стрелке», «Повернуть против часовой стрелки», «Повернуть на 180⁰», «Зеркально отразить относительно вертикали», «Зеркально отразить относительно горизонтали». На практиці звичайно вистачає перших двох можливостей.

КОРИГУВАННЯ РОЗДІЛЬНОЇ ЗДАТНОСТІ. Якщо зображення скановано з «неправильним» розділенням, зверніться до меню «Изображение → Исправить разрешение» або «Ctrl+Shift+T». У вікні встановіть потрібні параметри. Звичайне значення за замовчуванням «300dpi» змінювати не слід. Якщо всі сторінки скановані з низьким розділенням, переустановіть перемикач «Исправить разрешение» у положення «Всех изображений пакета».

ОЧИЩЕННЯ ЗОБРАЖЕНЬ ВІД ДЕФЕКТІВ. Тепер спробуємо поліпшити якість картинки. Відразу скажемо: серйозні дефекти зображення Finereader не виправить. Він може прибрати лише незначні дефекти, такі як смітинки на папері, невеличкі плями й т. і. Причому цю операцію він зробить інтелектуально: текст залишиться у первісному стані, його пошкоджено не буде. Щоб поліпшити зображення і прибрати з нього незначні дефекти треба діяти так:

- виділити одну або кілька сторінок документа;

- викликати меню «Изображение → Обработать изображение → Очистить изображение от мусора».

«ВИРІВНЮВАННЯ» РЯДКІВ. Один з типових дефектів сканування – це деформація рядків тексту. Ось приклад: при скануванні книги текст, що міститься біля її переплету звичайно виглядатиме нерівним. У цьому випадку рядки тексту на сканованому зображенні будуть виглядати не прямими, а трохи деформованими. Виправити такий дефект може операція усунення деформації рядків. Правда, упоратися із цим завданням вона зможе не завжди!

Щоб вирівняти рядки у відсканованому зображенні зробіть так.

- виділіть одну або кілька сторінок документа;

- зверніться до меню «Изображение → Обработать изображение → Устранить искажение строк».

Користуйтеся інструментом вирівнювання рядків обережно. У деяких випадках він може відробити неправильно. Тому не застосовуйте вирівнювання відразу до всіх сторінок документа. Обробляйте їх по однієї, щораз переглядаючи отриманий результат. Цими можливостями ми поки що й обмежимося.

4.5 Розмітка документа

Інформацію в електронному документі можна умовно розділити на три категорії: текст, таблиця й рисунок. При розпізнаванні програмі FineReader бажано «знати», з чим вона має справу. У принципі, аналізуючи сторінку, виявити її структуру FineReader може й самостійно. Але при складному форматуванні можливі помилки. Наприклад, якщо в документ вставлене зображення вікна Windows. FineReader напевно спробує розпізнати текст меню цього вікна, хоча за змістом нічого такого робити не потрібно.

У простих випадках опирайтеся на автоматичну розмітку тексту. Якщо форматування йде в декількох колонках, перемежується малюнками, таблицями – робіть розмітку самі. виправляти помилки FineReader-а вийде довше. Тому треба освоїти «ручну» розмітку документа, тим більше, що зробити це зовсім не складно.

Подивіться на панель інструментів розмітки ліворуч від області сторінок. Почнемо з іконки з буквою «Т»: це інструмент для визначення текстового блоку. Клацніть на іконці лівою кнопкою миші. Поставте покажчик на зображенні документа там, де повинен починатися фрагмент тексту. Утримуючи натиснутої ліву кнопку миші, обведіть текстовий блок. Відпустите ліву кнопку миші. У документі з'явилася прямокутна область зеленого кольору. Вона повинна охоплювати фрагмент зображення, що FineReader тепер буде розглядати як текст. На рисунку 2 текстовий блок охоплює всю частину документа, за винятком таблиці. Він позначений номером 1.

Щоб видалити блок розмітки потрібно виділити його щигликом миші й нажати клавішу «Del». Щоб змінити розмір блоку виділіть його й розтягніть за кут або границю так, як ви міняєте розмір вікна Windows.

Таким же чином ви можете виділити блок з таблицею або рисунком. Ці іконки розташовані нижче інструмента виділення текстового блоку. З областями рисунків все зрозуміло: ці ділянки зображення FineReader розпізнавати не буде. Він перенесе їх до результуючого документу «як є». А от

з таблицями має сенс познайомитись докладніше. Зазвичай розпізнавання таблиці робиться так.

Вибираємо інструмент визначення блоку з таблицею.

Акуратно обводимо фрагмент зображення, де міститься таблиця. Зайвий простір при цьому краще не чіпати, але й зовнішня границя таблиці повинна повністю перебувати всередині блоку.

Клацаємо усередині створеного блоку правою кнопкою миші. З контекстного меню (рис. 3) вибираємо «Анализ структуры таблицы».

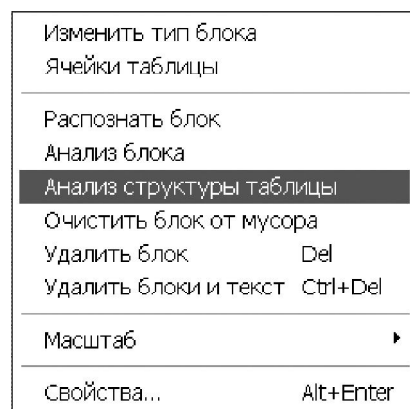


Рисунок 3

Переглядаємо результат розмітки. Особливу увагу звертаємо на те, як розташувались окремі осередки таблиці. Для цього є сенс послідовно клацати лівою кнопкою миші по осередках таблиці в області розмітки й аналізувати правильність їхнього розташування.

Якщо FineReader усе зробив правильно, переходимо до наступного етапу роботи.

А що робити в тому випадку, якщо при розмітці таблиці є певні помилки? Наприклад, FineReader не розпізнав бліду лінію границі й об'єднав декілька осередків в один? Для цього на панелі розмітки блоків є інструменти: «Добавить вертикальную линию», «Добавить горизонтальную линию» й «Удалить линии». Користуватися ними дуже просто. Клацніть на іконці «Добавить вертикальную линию». Переведіть покажчик миші у віконце «Изображение» на таблиці. Разом з покажчиком по зображенню буде переміщатися вертикальна лінія розмітки. Поставте її на потрібне місце й клацніть лівою кнопкою миші. Таким же чином додайте всі відсутні вертикальні лінії в розмітці таблиці. Щоб закінчити роботу в режимі додавання ліній виберіть інший інструмент розмітки. Таким же чином розберіться із горизонтальними лініями. Інструмент «Удалить линии» допоможе стерти зайву розмітку таблиці.

При роботі з таблицями часто доводиться поєднувати або розділяти окремі осередки таблиці. Зробити це простіше всього через праву кнопку миші. Поставте покажчик на осередок таблиці. Клацніть правою кнопкою миші. З контекстного меню (рис. 3) виберіть пункт «Ячейки таблицы». У ньому є три

розділи: «Разбить ячейки», «Объединить ячейки», «Объединить строки». Якщо FineReader помилково злив кілька осередків в одну, скористайтеся з можливості «Разбить ячейки». Якщо буде потрібно об'єднання осередків в одне ціле, виділіть їх і зверніться до пункту «Объединить ячейки» контекстного меню.

4.6 Розпізнаємо зображення

Коли розмітка сторінок завершена, натисніть «Ctrl+Shift+R» – розпізнати всі сторінки. Цю дію можна виконати через меню «Процесс → Распознать → Распознать все» або розкривши список доступних дій кнопки «Распознать».

Стежте за тим, яка мова встановлена для розпізнавання. Вибрати мову можна прямо на панелі інструментів FineReader. Наприклад, на рис. 2 встановлена пара «Русский и английский». Щигликом миші розкрийте список і виберіть варіант, що підходить для вашого випадку.

Результат розпізнавання можна подивитися у віконці «Текст» (рис. 2). Уважно прогляньте його. Спочатку переконаєтеся, що FineReader правильно обробив структуру документа. Якщо це не так, відкоригуйте розмітку й заново розпізнайте сторінку.

При повторному розпізнаванні не обов'язково заново обробляти весь документ. Ви можете обмежитися поточною або декількома виділеними сторінками. Виділіть сторінку й клацніть на кнопці «Распознать» або натисніть «Ctrl+R».

4.7 Зберігаємо результат

Робота над великими документами може забрати багато часу й проходити в декілька сеансів. Щоб зберегти результат роботи в програмі FineReader використовується поняття пакет. Пакет поєднує вихідні зображення, їхню розмітку, результати обробки в тім стані, що було на момент створення пакета. Зберегти пакет можна через меню «Файл → Сохранить пакет как...». З'явиться вікно, ідентичне вікну Провідника. Вибираємо папку, вказуємо ім'я пакета й результат збережений.

Пакет являє собою безліч файлів. Щоб не вносити сум'яття, не зберігайте пакет у папку, де вже є якісь дані. Краще завжди зберігати пакет в окремому каталозі.

4.8 Імпортуємо результат

Імпорт потрібний для збереження результатів у зовнішньому додатку або буфері обміну. Щоб зберегти поточну сторінку в Word клацніть на останній іконці з номером 4 і написом «Microsoft Word» (рис. 2). Відкриється вікно Word з розпізнаним документом. Це не єдиний спосіб імпортування даних. Клацніть на значку розкриття списку кнопки «Microsoft Word». З'являться такі варіанти: «Мастер сохранения результатов», «Сохранить страницы...», «Передать страницы в...», «Передать все страницы в...», «Опции».

Через «Сохранить страницы...» можна записати результат розпізнавання на диск. При записі є можливість вибрати всі або поточну сторінки, а також визначити формат файлу для збереження.

Звичайно користуються пунктами «Передать страницы в...» і «Передать все страницы в...». Ці дві можливості відрізняються тільки тим, що в першому випадку імпортується поточна або кілька виділених сторінок, у другому випадку FineReader передасть всі сторінки документа. Обоє меню містять кілька розділів. Їхня кількість залежить від установлених програм на вашому комп'ютері. Можна обмежитися трьома: «Microsoft Word», «Microsoft Excel» й «Буфер обмена». Виберіть один із зазначених пунктів. FineReader викличе додаток і передасть у нього результат розпізнавання.

4.9 Розпізнавання таблиці складної структури

Таблиці неоднорідної структури на практиці зустрічаються досить часто. Тому, завершуючи тему розпізнавання документів, дамо пораду, як з ними працювати.

Перш ніж розпізнавати складну таблицю постарайтеся визначити її «базову структуру». Визначити той «оптимальний» набір рядків і колонок, що згодом мінімізує витрати на поділ й об'єднання осередків. Коли структура проясниться, зробіть ручну розмітку таблиці.

Вибір структури дуже важливий. Не починайте обробку документа, поки не опрацюєте це питання. Погляньте, наприклад, на бланк податкової накладної, зображений на рисунку 4. Спроба розпізнати документ в автоматичному режимі, скоріш за все, не вдасться. Буде отримано набір неупорядкованих в текстових полів, фрагментів таблиць тощо. Надати лад такій абракадабрі буде складніше, ніж надрукувати бланк «з нуля». Проблема вирішується просто. Більшість частин документу складається з 11 колонок. Вони пронумеровані й показані у верхній частині малюнка.

1	2	3	4	5	6	7	8	9	10	11
Оригінал (видається покупцю)					ЗАТВЕРДЖЕНО наказом ДПА України від 30.05.97 р. N 165(у редакції наказу ДПА України від 08.10.98 р. N 469)					
Копія (залишається у продавця)										
(Непотрібне виділити поміткою "X")										
ПОДАТКОВА НАКЛАДНА										
Дата виписки податкової накладної				05.05.07		Порядковий номер			505	
Продавець							Покупець			

Розділ	Дата відвантаження (виконання, поставки (оплати) товарів (послуг) продавця)	Номенклатура поставки товарів (послуг) продавця	Одиниця виміру товару	Кількість (об'єм, обсяг)	Ціна поставки одиниці продукції без урахування ПДВ	Обсяги поставки (база оподаткування) без урахування ПДВ, що підлягають оподаткуванню за ставками				Загальна сума коштів, що підлягає оплаті
						20 %	0 % (реалізація на митній території України)	0 % (експорт)	Звільнення від ПДВ	
1	2	3	4	5	6	7	8	9	10	11
I	05.05.07	Маркер	шт.	10	1,5	15,00				
		Ручка	шт.	9	1,8	16,20				
		Ластик	шт.	4	0,65	2,60				
		Усього по розділу I				33,80				33,80
II		Товаротransпортні								1,80

Рисунок 4

При розмітці виділили блок типу «Таблиця». Інструментами «Додать вертикальну лінію», «Додать горизонтальну лінію» розмітили таблицю. Зазначте, ми це зробили вручну. Після цього пройшлися по таблиці, поєднуючи окремі осередки заголовка в групи. Результат розпізнавання такого складного документа був чудовий. Нагадаємо, що поєднувати й розділяти осередки зручно через праву кнопку миші (пункти «Разбить ячейки», «Объединить ячейки»). Якщо ви вдало вибрали структуру для ручної розмітки, найбільш складні таблиці вам не страшні.

5 Коригування зображень перед їхнім розпізнаванням

5.1 Як перезняти зображення

Якщо ви користуєтеся сканером, сміло пропустіть цей розділ або прочитайте його «по діагоналі». Матеріал цей для тих, хто збирається «сканувати» документи цифровим фотоапаратом або камерою мобільного телефону. От кілька простих порад по організації процесу.

1. Ахіллесова п'ята будь-якого фотоапарата – чутливість до висвітлення. Чим більше буде світла, тим кращим буде результат. Знайдіть можливість добре освітити документ, перш ніж скопіювати його.

2. Добийтеся рівномірного висвітлення сторінки документа з усіх боків. Іноді дзеркальце з протилежної сторони від джерела світла допомагає вирішити проблему.

3. Стежте, щоб тінь від вашого тіла (або його частини) не падала на документ. Забрати темну пляму на аркуші буде непросто. Навчіться критично оцінювати предмет, що фотографуєте.

4. Стежте за кутом нахилу камери (телефону). Типова помилка при перефотографуванні документів – зйомка під неправильним кутом. У цьому випадку геометрію зображення буде порушено. Замість прямокутного аркуша ви одержите трапецію. Усунути цей дефект можна. Але краще тримати камеру перпендикулярно до площини документу.

5.2 Штучний спалах та його використання

1. Включіть спалах на камері, якщо він є. Це додасть світла, зображення буде чітким й контрастним. Але ж головне – всі зображення будуть приблизно однакові. Коригувати їх можна буде за однією схемою. А це спрощує завдання.

2. При роботі зі спалахом не знімайте з близької відстані. Швидше за все, спалах «пересвітить» кадр, документ вийде занадто світлим. Можна перезнімати з відстані 70–90 см. Точне значення вам підкаже досвід спілкування з вашою моделлю фотокамери. Щоб оригінал не вийшов дуже дрібним, виберіть ступінь збільшення (зуммування) у 1,5, максимум в 2 рази. Тоді і спалах кадр не «пересвітить», й погіршення геометрії будуть мінімальні.

3. Максимально використайте площину кадру. Стежте, щоб не було великих полів між краями аркуша й границями зображення. До речі, за цим критерієм можна оцінити ступінь зуммування.

4. Не встановлюйте граничне збільшення: швидше за все, виникнуть проблеми з різкістю зображення.

І наостаннє. Запам'ятаєте, що від якості зйомки залежить обсяг роботи з коректування зображення. Чим акуратніше ви перезняли оригінал, менше прийде його коректувати. Витративши лічені секунди на поліпшення умов висвітлення, ви можете заощадити години роботи над поліпшенням фотокопії. І наостаннє, давайте зупинимось на тих проблемах, що виникають незалежно від технічних питань. Це, в першу чергу, порушення геометрії, світла й тону.

5.3 Послідовність дій з коригування зображень

Отже документи зняті, файли переписані до жорсткого диску. Перш ніж приступити до розпізнавання, потрібно подивитися, що ми маємо у розпорядженні. А одержали ми, скоріш за все, зображення як на рисунку 5.

Це зображення одержано за допомогою камери мобільного телефону з розділенням 2 Мегапіксела в приміщенні. Звичайно, високої якості очікувати не доводиться. На зображенні є всі типові проблеми для цього випадку. Центр і краї засвічені нерівномірно. Папір виглядає сірим, у той час як реально він білого кольору. Є геометричні перекручування, текст на сторінці має трапецієподібну форму. Подивимось, що можна зробити. Для роботи нам знадобиться редактор растрової графіки. Зупинимо свій вибір на PhotoShop CSx.

Мабуть, це кращий професійний редактор. PhotoShop – різноманітна й складна програма. Про її можливості ходять легенди. Ми не збираємось докладно їх вивчати. Нам потрібний тільки результат. Тому основні прийоми коректування будемо розглядати без докладних коментарів. Приймайте їх як є і дійте за принципом «роби як я».

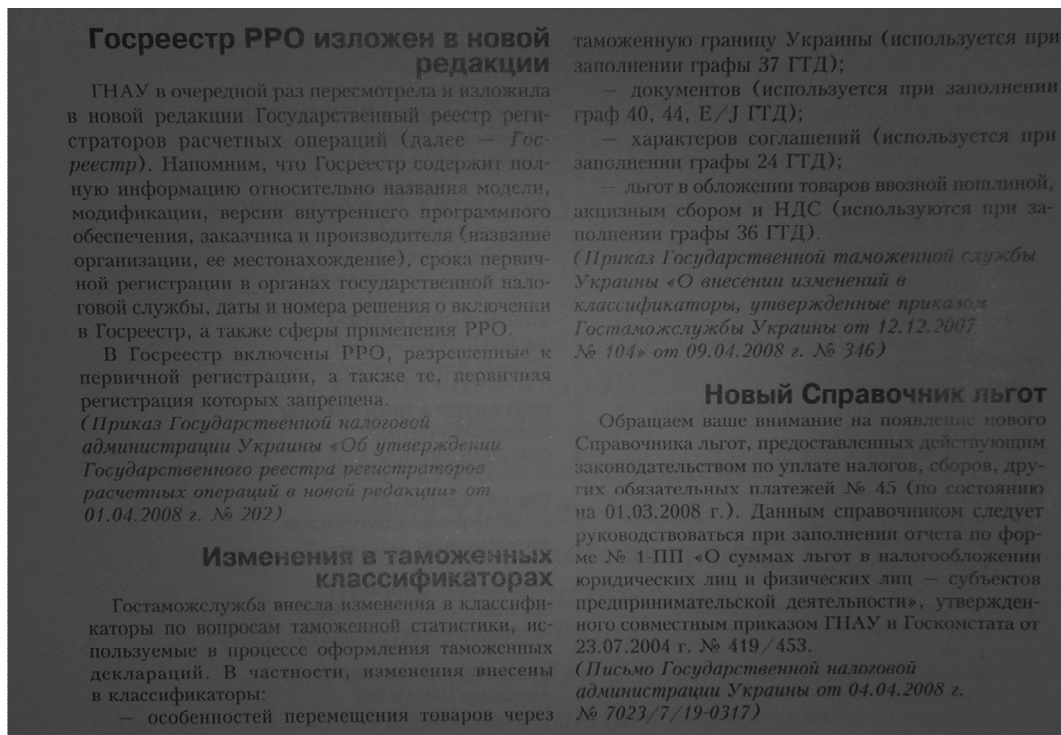


Рисунок 5

5.3.1 Завантажуємо файл

Викликаємо PhotoShop через «Пуск → Програми». Відкриваємо в ньому зображення. Для цього тиснемо «Ctrl+O», з'явиться вікно Провідника. У ньому знаходимо файл подвійним щигликом відкриваємо в програмі. Вікно PhotoShop із завантаженим зображенням показане на рисунку 6.

5.3.2 Коригуємо розмір зображення

Це перше, з чого ми почнемо. І ось чому. Матриця камери була 2 Мегапікселя. Це означає, що оригінал, займаючи весь кадр, буде розміром приблизно 9x12 см у розділенні 300 dpi. Тепер має сенс повернути розміри до первісних: великі букви FineReader розпізнає краще.

Натискаємо «Alt+Ctrl+I» («Изображение → Размер изображения»). З'явиться вікно, показане на рисунку 7. Вимірюємо (приблизно) розмір оригіналу. Досить визначити один вимір, наприклад – ширину. Вводимо отримане значення в поле «Ширина».

Прапорці «Інтерполяція» й «Зберегти пропорції» повинні бути включені. Значення параметра «Дозвіл» повинне бути не менш «300 dpi».

У нашому випадку ширина зображення була близько 21 см. Вказали у поле «Ширина» значення «21» і натиснули «Ок». Якщо ви все зробили правильно, розмір зображення (у мегабайтах) повинен змінитися.

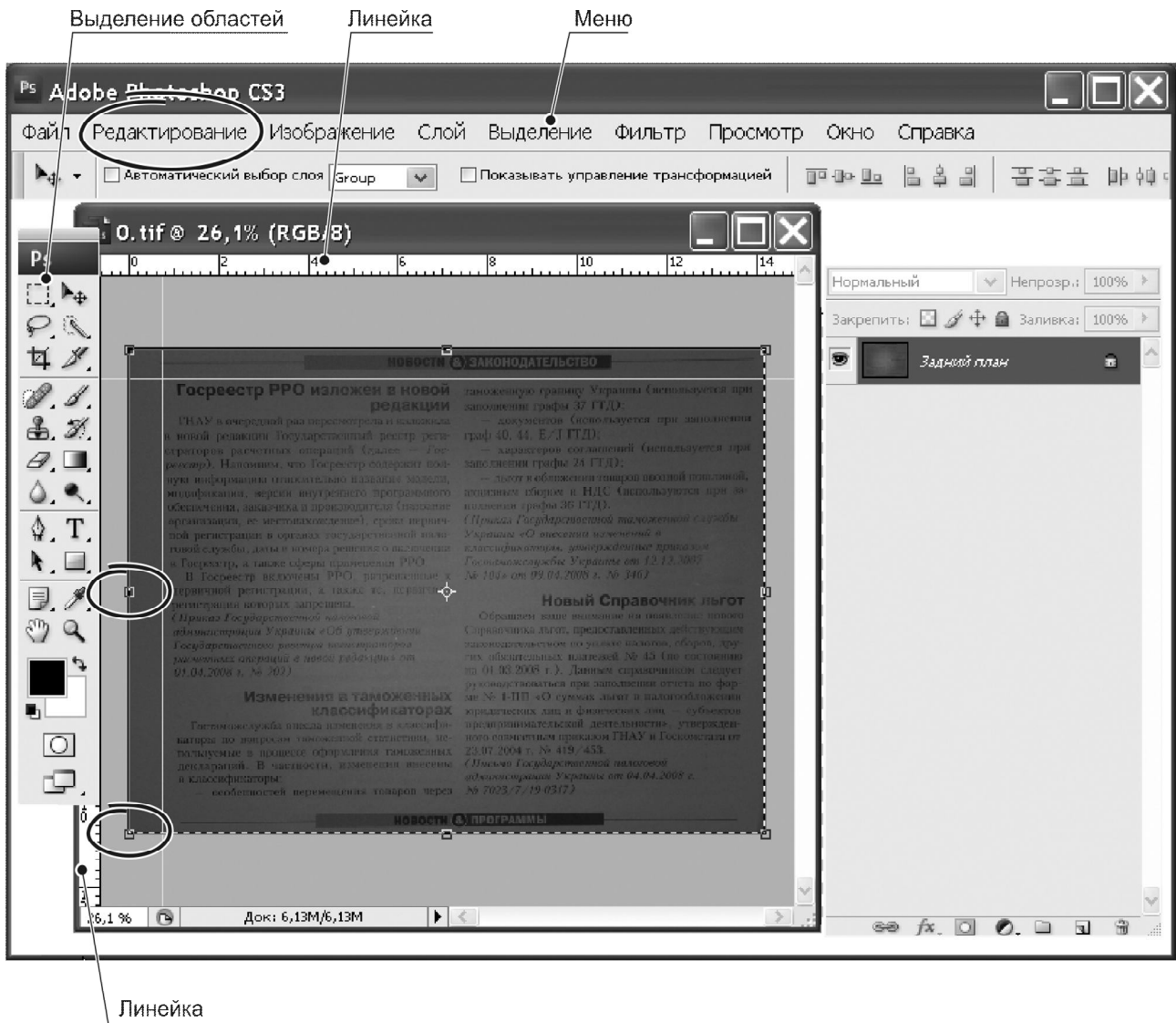


Рисунок 6

5.3.3 Прибираємо зайві деталі

1. Натискаємо «М», букву «М» натискайте на англійській клавіатурі. Ми вибрали інструмент «Прямоугольное выделение».

2. Утримуючи ліву кнопку миші, обведіть площу на зображенні.

3. Відпустіть кнопку миші, навколо області з'явиться пунктирна рамка.

4. Викличте меню «Изображение → Кадрировать». Photoshop прибере фрагменти документа, що розташовані поза виділеною областю.

5.3.4 Коригуємо геометрію

1. Поставте покажчик миші на ліву лінійку вікна із зображенням (рис. 6).

2. Утримуючи ліву кнопку миші, перемістіть її покажчик на область документа. З'явиться вертикальна напрямна.

3. Повторіть операцію, перетягнувши напрямну з верхньої лінійки на документ. Так ми отримали орієнтири для контролю перпендикулярності.

4. Натискаємо «Ctrl+A» (символ «А» тиснемо на англійській клавіатурі). Ми виділили весь документ. Довкола нього з'явиться пунктирна рамка.

5. Натискаємо «Ctrl+T» («Редактирование → Свободная трансформация»). Навколо зображення з'явиться прямокутна область із прямокутними мітками на краях і на кожній стороні. Розтягуючи область, ми збільшуємо розмір зображення в тому або іншому напрямку. Але нам потрібно інше: «розтягти» нижню частину зображення вправо. Ставимо покажчик миші на прямокутник у правому нижньому куті. Натискаємо «Ctrl» і тягнемо його вправо, поки текст не вирівняється відносно сторінки.

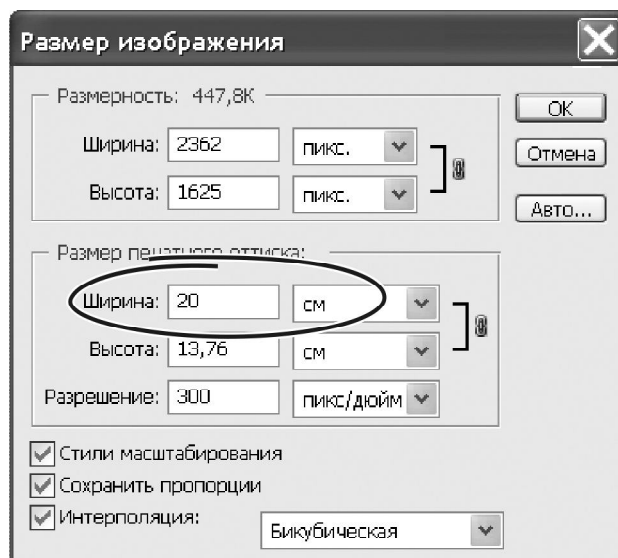


Рисунок 7

5.3.5 Коригуємо яскравість

1. Натискаємо «Shift+D», щоб зняти виділення (про всякий випадок).

2. Дивимося на значок виділення областей (рис. 6). Якщо він прямокутний, то тиснемо «М», потім «Shift+M». Значок виділення стане овальним. Якщо значок був овальний, просто натискаємо «М».

3. Утримуючи натиснутої ліву кнопку миші, обводимо більше яскравий фрагмент зображення (у центрі). Відпускаємо ліву кнопку миші.

4. Натискаємо «Q». Зображення покриється червоною «поволокою». Залишиться тільки «отвір» у центрі» із чіткими краями.

5. Натискаємо «Фільтр → Размытие → Размытие по Гауссу». З'явиться вікно з єдиним параметром. «Радіус». Збільшуйте цей параметр, щоб домогтися дуже плавного розмиття країв «отвору». Зазвичай радіус становить 150-200.

6. Натисніть «Ок», щоб закрити вікно.

7. Натисніть «Q». Червона «маска» пропаде (ми перейшли в режим роботи із зображенням) .

8. Натисніть «Ctrl+I» (інвертувати виділену область).

9. Натисніть «Ctrl+L». З'явиться вікно, зображене на рисунку 8. Ставимо покажчик миші на центрального повзунка й, утримуючи натиснутої ліву кнопку, плавно переміщаємо його вправо. Краї зображення повинні стати світлішими. Описаний прийом потребує певного досвіду. Від того, наскільки вдало побудована виділена область, як плавно вона розмита залежить результат вирівнювання. Ви можете повторити процес з іншим розміром виділеної області. Але кожного разу описаний спосіб дозволяє одержати прийнятний результат.

10. Натискаємо «Ок».

11. Знімаємо виділену область «Ctrl+D».

12. Знову вибираємо «Ctrl+L».

13. Щоб висвітлити зображення, переміщаємо правого повзунка ліворуч. Щоб підвищити щільність тексту, переміщаємо лівого повзунка йому праворуч. Не прагніть одержати текст на ідеально білому фоні. Текст на сіруватому фоні FineReader теж розпізнає.

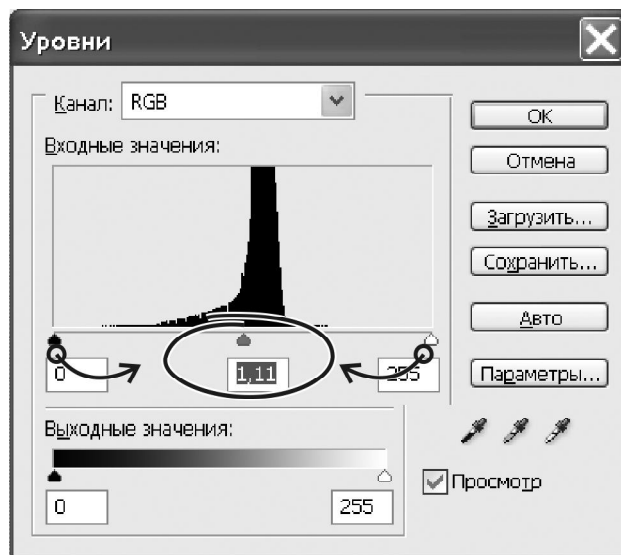


Рисунок 8

14. Щоб закрити вікно, натискаємо «Ок».

15. Зберігаємо документ «Shift+Ctrl+S» («Сохранить как...») або «Ctrl+S» (просто «Сохранить»). На рисунку 9 показано результат нашої роботи.

Зверніть увагу, – цей результат було отримано за допомогою мобільного телефону з камерою низької якості. Залишається тільки розпізнати файл.

6 Завдання до виконання роботи

Вхідну інформацію для роботи студент обирає сам. Це має бути текстовий документ (або декілька різних документів) на паперовому носії. Метою роботи буде розпізнавання такого документу. У документі для розпізнавання має бути не менше 5 таблиць та 5 рисунків. Причому рисунки мають бути розташовані за текстом. Розпізнавання сторінок, де розташовано один рисунок, не розглядається. У звіті щодо виконання роботи студент має надати детальне описання кожного кроку розпізнавання документа. Кожний етап роботи треба проілюструвати відповідним рисунком (скриншотом). Звіт опублікуйте на курсі

Госреестр РРО изложен в новой редакции

ГНАУ в очередной раз пересмотрела и изложила в новой редакции Государственный реестр регистраторов расчетных операций (далее — *Госреестр*). Напомним, что Госреестр содержит полную информацию относительно названия модели, модификации, версии внутреннего программного обеспечения, заказчика и производителя (название организации, ее местонахождение), срока первичной регистрации в органах государственной налоговой службы, даты и номера решения о включении в Госреестр, а также сферы применения РРО.

В Госреестр включены РРО, разрешенные к первичной регистрации, а также те, первичная регистрация которых запрещена.

(Приказ Государственной налоговой администрации Украины «Об утверждении Государственного реестра регистраторов расчетных операций в новой редакции» от 01.04.2008 г. № 202)

Рисунок 9

<http://cdo.kname.edu.ua/course/view.php?id=285>.

Список рекомендованої літератури

1. Карпенко Н. Ю. OpenOffice для бухгалтера, аудитора, економіста / Н. Ю. Карпенко, Ю. В. Цыганок. – Харків: ИКК Фактор, 2000. – 496 с.
2. Конспект лекцій з курсу «Інформатика і системологія» (для студентів 2 курсу заочної форми навчання напряму підготовки 6.040106 «Екологія, охорона навколишнього середовища та збалансоване природокористування» / М. Ю. Карпенко, В. Б. Уфимцева, О. В. Димченко, Н. К. Корзун. – Харків: ХНУМГ, 2012. – 123 с. – (поз. 154Л)
3. Карпенко М. Ю. Конспект лекцій з курсу «Управлінські інформаційні системи в аналізі і аудиті» (для студентів 5 курсу всіх форм навчання за напрямом підготовки 6.030509 «Облік і аудит») / М. Ю. Карпенко, В. Б. Уфимцева. – Харків: ХНУМГ, 2012. – 93 с. – (поз. 156Л)

Виробничо-практичне видання

Методичні рекомендації

до самостійної роботи з дисциплін
**«Електронний документообіг в проектувальній діяльності»,
«Інформаційні технології в науковій та презентаційній роботі»,
«Інформаційні системи в менеджменті»**

(для студентів усіх форм навчання спеціальності 073 – Менеджмент)

Укладачі: **КАРПЕНКО** Микола Юрійович,
ГАВРИЛЕНКО Ірина Олександрівна

Відповідальний за випуск *О. Б. Костенко*

За авторською редакцією

Комп'ютерне верстання *М. Ю. Карпенко*

План 2017, поз. 418М

Підп. до друку 24.02.2017 Формат 60х84/16
Друк на ризографі Ум. друк. арк. 1,0
Зам. № Тираж 50 пр.

Видавець і виготовлювач:

Харківський національний університет
міського господарства імені О. М. Бекетова,
вул. Маршала Бажанова, 17, Харків, 61002
Електронна адреса: rectorat@kname.edu.ua
Свідоцтво суб'єкта видавничої справи:
ДК № 5328 від 11.04.2017.